# Chiral effects on helicity studied via the energy landscape of short (d, l)-alanine peptides

Neelamraju, Sridhar; Oakley, Mark T.; Johnston, Roy L.

*DOI:*
10.1063/1.4933428



*Document Version*
Publisher's PDF, also known as Version of record

# Chiral effects on helicity studied via the energy landscape of short (d, l)-alanine peptides

Sridhar Neelamraju, Mark T. Oakley, and Roy L. Johnston

Citation: The Journal of Chemical Physics **143**, 165103 (2015); doi: 10.1063/1.4933428
View online: http://dx.doi.org/10.1063/1.4933428
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/143/16?ver=pdfcov
Published by the AIP Publishing

# Chiral effects on helicity studied via the energy landscape of short (D, L)-alanine peptides

Sridhar Neelamraju,[a] Mark T. Oakley, and Roy L. Johnston
*School of Chemistry, University of Birmingham, Edgbaston B15 2TT, United Kingdom*

(Received 26 June 2015; accepted 7 October 2015; published online 28 October 2015)

The homochirality of natural amino acids facilitates the formation of regular secondary structures such as $\alpha$-helices and $\beta$-sheets. Here, we study the relationship between chirality and backbone structure for the example of hexa-alanine. The most stable stereoisomers are identified through global optimisation. Further, the energy landscape, a database of connected low-energy local minima and transition points, is constructed for various neutral and zwitterionic stereoisomers of hexa-alanine. Three order parameters for partial helicity are applied and metric disconnectivity graphs are presented with partial helicity as a metric. We also apply the Zimm-Bragg model to derive average partial helicities for Ace-(L-Ala)$_6$-NHMe, Ace-(D-Ala-L-Ala)$_3$-NHMe, and Ace-(L-Ala)$_3$-(D-Ala)$_3$-NHMe from the database of local minima and compare with previous studies. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4933428]

## INTRODUCTION

The thermodynamic drive to form stable $\alpha$-helices that fold quickly could explain why our protein alphabet is predominantly homo-chiral as was first suggested by Pauling and Corey.[1,2] For most naturally occurring amino acids, other than glycine, the chirality of the $\alpha$-carbon atom directly influences secondary and tertiary structures of peptides and proteins. An $\alpha$-helix composed of L-amino acids will usually be right-handed, while one composed of D-amino acids will usually be left-handed.[3] Occasionally, depending on the side-chain, some $\alpha$-helices with L-amino acids can be left-handed as well. Ooi *et al.*[4] postulated that non-bonded interaction energy favours the left-handed alpha helix over the right-handed and the dipole-dipole interaction between the side-chain and the backbone can be sufficiently large to reverse the screw-sense of the helix. Homochirality of the natural amino acids facilitates the formation of regular secondary structures such as $\alpha$-helices and $\beta$-sheets. Thus, altering the stereochemistry of one or more $\alpha$-carbons has a profound impact on the structure and stability of a given peptide.

Most naturally occurring peptides are composed entirely of L-amino acids. However, many examples exist in nature where polypeptides contain both L- and D-amino acids.[5,6] Altering the stereochemistry of the $\alpha$-carbons can lead to novel backbone structures. For example, a bracelet-shaped backbone was synthesized in a fourteen-residue peptide,[7,8] while a boat-shaped molecular fold was synthesized in another twenty residue peptide[7] through the switching of strategic amino acid residues from the L form to the D form. The cyclic decapeptide, gramicidin S, was synthesized by Hodgkin *et al.*[9] where two of the ten residues were D-Phe. We study here how the stereochemistry of the polypeptide backbone influences the stability and structure of small alanine polymers. Further, we explore the possibility of locating helical and partially helical configurations among various stereoisomers of alanine hexamers by performing global and local optimizations and constructing detailed energy landscapes. Here, we strive to locate within the potential energy landscape partially helical segments and determine their correlation to the energies of local minima and transition states.

We describe four order parameters designed to measure the helicity of the peptide conformers. In the past, theoretical studies of the stereochemistry of peptides have focused on hydrogen bonding patterns and torsional angles ($[(\phi,\psi) = (-65 \pm 35, -37 \pm 30)]$) for three intervening residues to be in the helical arrangement. For short peptides of up to six residues, we find that this criterion is not sufficient to capture the nature of backbone structure (especially helicity of a single residue) since one needs a minimum of three residues to fold into an alpha-helix to use the hydrogen bonding criterion.[10–12] In our simulations, we often find structures with one to three residues folding into an alpha-helix and the rest into a random coil. Further, the use of coarse-grained potentials that predefine the torsion angles between radii also biases the simulations in favour of helical backbones.[7,11,13,14] Helix-coil transition theory models are better supported with explicit definitions of whether a given amino acid residue is in a helical configuration or not.[15] Many studies have been performed on ascertaining quantitatively, purely from the geometry, the number of consecutive residues in an $\alpha$-helical state.[16–20]

One way to visualize and summarise the connections between all the local minima and transition states found is the disconnectivity graph.[21,22] We use our helicity measures to produce metric disconnectivity graphs[23] to present both the energy landscape and the helicity simultaneously.

## METHODS

### Mapping the energy landscape

We investigate the energy landscapes of alanine oligomers with 4-6 residues. For each oligoalanine, we consider

[a]Author to whom correspondence should be addressed. Electronic mail: s.neelamraju@bham.ac.uk

 

all combinations of D- and L-alanine residues. We generate the initial structures using the tleap program available in the AmberTools package.[24] The energies of the peptide conformers are evaluated using the AMBERff03 force-field,[25] with solvation modelled using the generalised Born model.[26] We use the basin-hopping algorithm[27] implemented in GMIN to locate low-energy conformers of the peptides[28] Databases of minima and the transition states connecting them are then generated with discrete path sampling as implemented in PATHSAMPLE.[29] We select pairs of minima with the missing connection algorithm[30] and locate transition states between these pairs using the doubly nudged elastic band method,[31] interpolating between end points using a Cartesian coordinate interpolation scheme.[32] This is further optimized by hybrid eigenvector following in OPTIM.[33] In order to remove artificial frustration, the UNTRAP method was used for selecting pairs of minima for further nudged elastic band calculations. The whole networks of stationary points, consisting of the low-lying minima and transition states representing the system are visualised using disconnectivity graphs.[21,22] Order parameters are plotted on the disconnectivity graphs according to the helicity metric generated from the PyConnect program.[23]

## Helicity indices

Alanine hexamers display a wide range of structures. In the following, we describe five different methods to estimate partial helicity within a polymer from geometric considerations.

### *Projection perpendicular to helical axis*

In order to determine the helical nature of a structure, we perform the following analyses:

- Find the helical axis using the algorithm described by Aqvist.[34]
- Align the molecule parallel to the x-axis.
- Project the coordinates of the backbone of the peptide on the y-z plane.
- Fit these coordinates to a circle of radius "$R_{fit}$."

Using a combination of the radius and the standard deviation of the coordinates from the circle of radius $R_{fit}$, we characterize "helicity" in our short peptides. We note that there is a difference of about 5°-10° in the helical axis found by this method and the helical axis determined by using the difference in quarternions between frames comprising of consecutive residues as described below. The agreement becomes much better for larger peptides comprising eight or more residues. Of course, this method is useful for characterising completely helical structures and planar cyclic structures that are predominantly found for zwitterions.

### *Cluster compare algorithm*

As we discovered a number of structures in which one or more alanine residues were partially involved in the formation of a helix, we used the cluster compare algorithm[35] to analyse such structures. Here, we check for the existence of a subunit of a perfect right- or left-handed alpha helix comprising $n$ ($n$ = 1-6) alanine residues in all the structures found. We compare the backbone (...C–O–N–H–CA–HA–...) of a regular helix comprising n alanine residues with the backbone of the conformer under investigation. This method was originally developed for finding specific molecular motifs within bulk systems. Such an analysis allows for matching each alanine residue in the polypeptide with a reference "ideal" structure. In this case, the reference structures used are sections of a hexa-alanine residue in a left-handed $\alpha$-helical state. The number of residues in this reference structure is increased from 1 to 6 to ascertain the average helical length and average number of helices in the database of polypeptides.

### *Quaternion projection*

Another tool used for determining the helicity of the structure is the Screwfit method proposed by Kneller and Calligari.[36] Here, the helix parameters are derived from the superimposition of consecutive peptide planes. We can determine whether the superimposed residues are part of a helix or not. A residue is deemed to be helical if the helix radius (0.8-1.5 Å), angular distance (0.45-0.69 rad), error ($<10^{-8}$) in the superimposition fit, and the handedness of the helix are within the parameters pre-defined limits. Thus, for a hexapeptide, four superimposition matches with a radius of ≈1.8 Å (in water) with an angular distance (≈0.54 Å) imply that every alanine residue is in an $\alpha$-helical state.

### *Circular dichroism spectra*

Circular dichroism (CD) spectra of peptides are sensitive to their secondary structures. Peptides with $\alpha$-helical structures exhibit a characteristic signal, with a positive band at 190 nm and a double minimum at 208 and 220 nm. The intensity of the signal at 220 nm is proportional to the fractional helicity of a peptide.[37] The CD spectra of peptides can be calculated using the matrix method,[38–43] which treats a polypeptide as a collection of amide chromophores that interact electrostatically. This method has been shown to reproduce the ellipticity at 220 nm and is also fast enough to allow for the rapid calculation of the spectra of several thousand structures. Here, we use a model of the amide chromophore where the parameters of the $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions are fitted to CASSCF/CASPT2 calculations on $N$-methylacetamide.[44] The fractional helicity, $f_H$ is calculated as

$$f_H = \frac{[\theta]_{220}}{[\theta_{H\infty}]_{220}(1 - \frac{k}{N})}, \quad (1)$$

where $[\theta]_{220}$ is the ellipticity of the CD spectrum at 220 nm, $[\theta_{H\infty}]_{220}$ is the ellipticity of a hypothetical infinitely long helix, and $k$ is a constant to account for end effects. Here, we use $k = 3$ and $[\theta_{H\infty}]_{220} = -37\,000^\circ\,\mathrm{cm}^{-3}\,\mathrm{mol}^{-1}$. This method gives a continuous range of $f_H$, with negative values corresponding to left-handed helices.

### *Hydrogen bonding*

One can estimate fractional helicity by measuring $\frac{\langle n_h \rangle}{N_h}$ which is the average number of helical hydrogen bonds per

molecule which is given by[45]

$$\langle n_h \rangle = \frac{\partial \ln Z}{\partial \ln \omega}. \quad (2)$$

Here, $Z$ describes the partition function and $\omega$ is the helix nucleation parameter in the Zimm-Bragg model.

### Application of helix-coil models

The statistical mechanical models proposed by Lifson-Roig[46] (LR) and Zimm-Bragg (ZB) characterize the helix-coil transition in terms of a helix nucleation parameter and a helix propagation parameter.[45] The fractional helical content and helical segments can be related to the partition function to obtain these parameters in the LR model. Further details of the LR model can be found in Ref. 45. Using a combination of average fractional helicity and average length of a helix, we can estimate purely from a Boltzmann-weighted average from the complete ensemble of low-lying minima the helix nucleation (s in ZB and $\nu$ in LR) and helix propagation ($\sigma$ in ZB and $\omega$ in LR) parameters.[47,48] The accuracy of these parameters depends on the quality of sampling, accuracy of the index used for describing helicity, and the potential used for determining the energies. Previous molecular dynamic studies[47] have often used torsion angles between successive alanine residues to ascertain helicity. They have noted the appearance of artefacts and structures that are not completely helical using this approach. Here, the Boltzmann-weighted average fractional helicity calculated from an ensemble of local minima is compared to fractional helicities derived from the helix nucleation and propagation parameters described in the literature[47,49] for Ace-(L-Ala)$_6$-NHMe. Wójcik *et al.*[50] have demonstrated experimentally that for random copolymers, nearest-neighbour interactions contribute to the same extent in the helix and coil forms and therefore do not affect the intrinsic values of helix nucleation and propagation parameters. Further, Scheraga[51] have argued that the conformation of an amino acid residue in a polypeptide or protein is determined largely by the short-range interactions between a side-chain and the atoms of the backbone of the same amino acid residue. The long-range interactions are compensated by polymer-polymer and polymer-solvent interactions and the helix propensity of an amino acid is independent of its nearest-neighbours.

Within the one-sequence approximation,[52] denaturation is pictured as proceeding by unwinding from the ends. The partition function, here, is given by

$$Z(N) = 1 + \sigma \sum_{n=1}^{N-2} u^{N-n-2} \nu^2 w^n (N - n - 1), \quad (3)$$

where $N$ is the number of residues that can possibly take part in an $\alpha$-helix, $n$ is the number of residues in one allowed helical sequence,[45] and $w$ and $\nu$ are the propagation parameter and the equilibrium constant for the helical conformation in the random coil, respectively. More details can be found in Refs. 45 and 46. The fractional helicity within the one-sequence approximation is therefore calculated as

$$\langle N_H \rangle = \left(\frac{1}{N-2}\right) \frac{\partial \ln Z(N)}{\partial \ln s}. \quad (4)$$

## RESULTS

### Global optimisation with fixed chiralities

Global optimisations of the methyl-capped neutral alanine tetramer (Ace-(Ala)$_4$-NHMe), pentamer (Ace-(Ala)$_5$-NHMe), and hexamer (Ace-(Ala)$_6$-NHMe) were performed by flipping the chiralities on the alpha-carbon systematically from the all-L amino acid sequence (naturally occurring) to the all-D sequence. If the total number of alanine residues is given by $n_{\mathrm{Ala}}$, of which $n_D$ are in the D form, then the number of stereo-isomers that can exist is simply ${}^{n_{\mathrm{Ala}}}C_{n_D}$. We perform global optimisations by basin-hopping on all of these possible stereoisomers within the constraint of keeping the chiralities fixed during the global optimisation run.

The most stable stereoisomers for (Ala)$_{4-6}$ are enumerated in Table I and depicted in Figure 1. As expected, we see a symmetric behaviour from the C and N terminals of the polymers for the tetramer, pentamer, and hexamer (Figure 2). For example, in the neutral hexamer, isomers with 2L+4D and 4L+2D have identical energies. This behaviour does not appear to depend on chain length. We note that when, in a given peptide sequence, the number of D and L alanines is equal, the global minimum is not an alternating sequence. Sets of two D type $\alpha$-carbons sit together in the sequence. In the case of the neutral hexamer, the most stable stereoisomers always have

TABLE I. Most stable configurations found for all alanine polymers with different chiralities after global minimisations using the basin hopping scheme. Each run comprised 50 000 basin hopping steps followed by L-BFGS local optimization. High temperature MD steps were used to overcome energetic barriers from local minima during the global optimisation. Implicit solvent generalised Born approximation in water was used to model solvent effects.

| Total number of alanine residues | Number of D residues | Most stable isomer | Energy (kcal/mol) |
|---|---|---|---|
| Neutral | | | |
| 4 | 0 | LLLL | −40.538 |
| | 1 | LLLD | −40.798 |
| | 2 | DLLD | −39.686 |
| | 3 | DDDL | −40.798 |
| | 4 | DDDD | −40.538 |
| 5 | 0 | LLLLL | −43.068 |
| | 1 | LLLDL | −43.769 |
| | 2 | LLLDD | −43.456 |
| | 3 | DDDLL | −43.456 |
| | 4 | DDDDL | −43.870 |
| | 5 | DDDDD | −43.068 |
| 6 | 0 | LLLLLL | −46.356 |
| | 1 | LLLLDL | −46.401 |
| | 2 | LLLLDD | −45.361 |
| | 3 | DLLLDD | −46.057 |
| | 4 | DDDDLL | −46.361 |
| | 5 | DDDDLD | −46.401 |
| | 6 | DDDDDD | −46.356 |
| Zwitterions | | | |
| 6 | 0 | LLLLLL | −107.430 |
| 6 | 3 | DLDLDL | −107.645 |

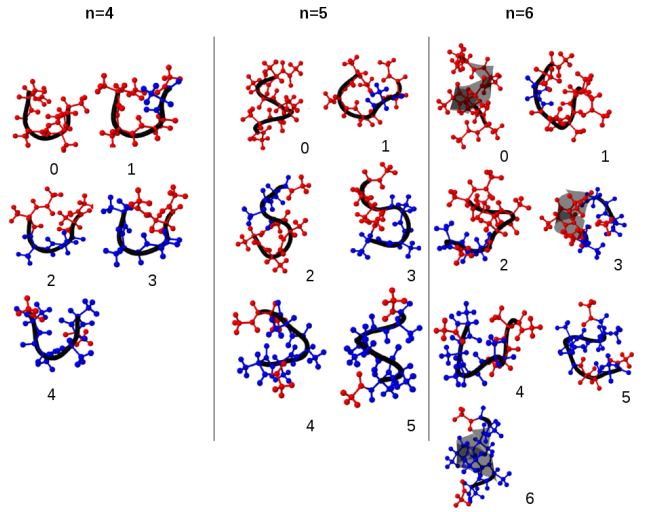


FIG. 1. Putative global minima obtained for the neutral alanine (Ace-$(Ala)_n$-NHMe), (n = 4, 5, 6) with chiralities systematically flipped on $\alpha$ carbons. Blue atoms indicate atoms belonging to the D-Ala and red atoms correspond to L-Ala. Capping groups are also in red. The shaded regions indicate the helical nature of the backbone along with the sense of the helix.

two D's or two L's sitting together in the chain. From lattice chain models of polyalanine, Nanda *et al.*[11] concluded that homochirality facilitates the formation of "maximally compact structures" and reduces the size of accessible fold space. Our all-atom models are in agreement with this observation.

The helix also allows for backbone inter-atomic contacts between $i$, $i+3$ and $i$, $i+4$ positions in sequence. In the ensemble of isotactic polyalanine hexamers, the lowest energy conformations are $\alpha$-helices. In LLLDDD and its sequence inverted stereoisomer (DDDLLL), the lowest energy conformations contain a pair of three-residue $\alpha$-turns.

## Characterising helicity on the energy landscape

Among amino acids, alanine has a relatively high helix propensity.[50,53] Polyalanines have been used as models to test various optimisation algorithms.[54] Energy landscapes are mapped for the following neutral isomers: (1) Ace-(L-Ala)$_6$-NHMe, (2) Ace-(L-Ala-D-Ala)$_3$-NHMe, (3) Ace-(L-Ala)$_3$-(D-Ala)$_3$-NHMe and two isomers with polar ends (zwitterions) that include (4) $^+NH_3$-(L-Ala)$_6$-$COO^-$ and (5) $^+NH_3$-(L-Ala-D-Ala)$_3$-$COO^-$. The effect of dielectric constant on helicity was also studied using the generalized Born implicit solvent model augmented with hydrophobic solvent accessible surface area as implemented within AMBER.[24]

The validity of the order parameters derived from the approaches mentioned above is first tested on Ace-(L-Ala)$_6$-NHMe. This system has been the subject of many global optimisation and energy landscape studies.[55] Linderstrøm-Lang[56] synthesised alanine with alternating D and L amino acids of around 30 residues that form an $\alpha$-helix. The possibility of partially folded left- and right-handed $\alpha$-helices was suggested by Lotan *et al.* for poly-(L-Ala-D-Ala-L-Ala-L-Ala)$_n$ (n = 1-6) sequences in water and 90% aqueous trifluoroethanol[57] and the average helical content was deduced from hypochromism measurements.

Nanda *et al.*[11] also studied the role of chirality in alanine hexamers (DLDLDL, DDDDDD, and LLLLLL) in shaping compactness and fold-space. In particular, they found that the configuration with alternating D and L alanines has a large ensemble of accessible configurations that are characterized by extended

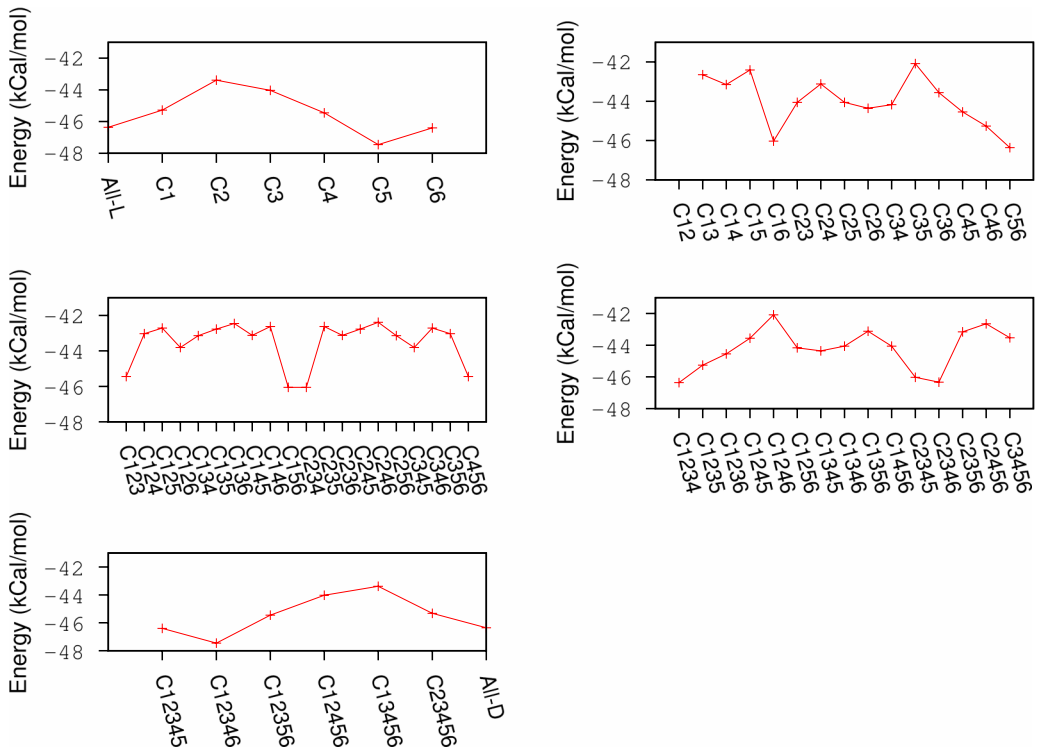


FIG. 2. Energies of putative global minima obtained for (Ace-(Ala)$_6$-NHMe) after basin-hopping runs with fixed chiralities for L-Ala. The x-axis indicates the number of the $\alpha$-carbons with the D-configuration. Numbering starts from the N-terminus.

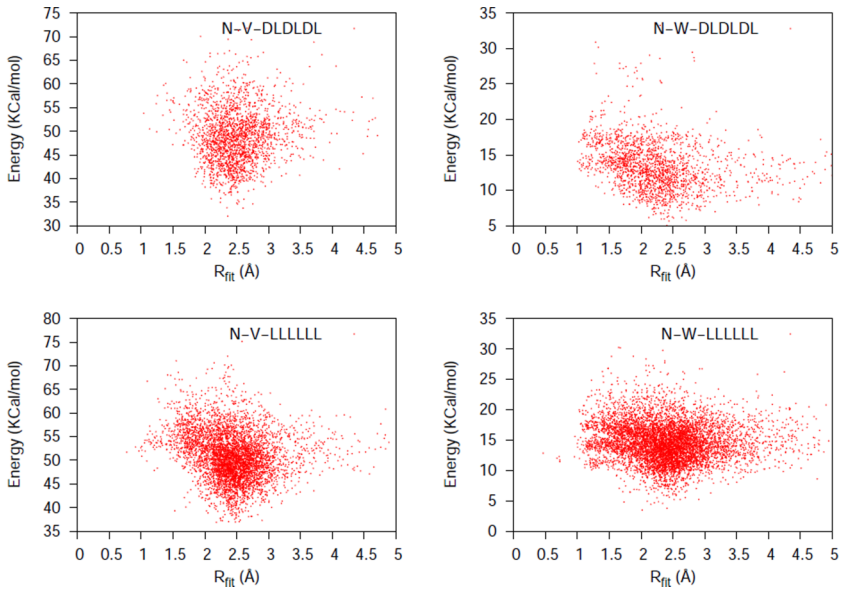


FIG. 3. Energy versus $R_{fit}$ parameter derived from the projection of the backbone along the plane perpendicular to the helical axis for $^{+}NH_3$-(L-Ala)$_6$-$COO^{-}$ in vacuum (bottom-left) and water (bottom-right) and $^{+}NH_3$-(L-Ala-D-Ala)$_3$-$COO^{-}$ in vacuum (top-left) and water (top-right).

structures as opposed to all L alanines that have easy access to folded states. Typically, the hydrogen bonding patterns and torsional angles are analysed to determine the nature of helicity in peptides.[12] However, this requires at least one turn of the helix to be complete. Many structures in our database display partial helicity with even two residues. It is easier to determine helicity from geometric information.

Figure 3 shows the $R_{fit}$ parameter as a function of energy for the LLLLLL and DLDLDL sequences in vacuum and water. We observe a clear clustering of low energy structures around the 2.5 Å radius. The helical system lies closer to the 2 Å radius in water and vacuum. Further, there are coiled structures in the database that tend to form ring-like structures that are almost planar. This is especially evident for the (L-Ala)$_6$ neutral isomer in vacuum (N-V-LLLLLL). The projection of the helical axis on the plane captures the cyclic nature of the backbone rather well. The lowest energy alpha-helix for the all-L neutral tetramer (optimised with AMBERff03 parameters) in water has a $R_{fit}$ value of 1.92 Å.

Partial helicity in peptides can be better quantified using the quaternion superimposition (QS) method as described in the "Methods" section.[36] We have constructed metric disconnectivity graphs for (L-Ala)$_6$. Here, the order parameter for helicity is the number of successful matches between superimposition of successive frames corresponding to a peptide residue. The major advantage of this method is that along with a superimposition fit, we can also calculate values of the radius of helix, direction of the helix, angular distance between successive frames, and a straightness parameter. We calculate these values for an optimised $\alpha$-helix and then use these set thresholds to accept or reject a given structure. Selection of stringent criterion on the basis of the known structure makes the characterisation of partial helicity more accurate. CD spectra can also provide an estimate of the fractional helicity of a peptide. Disconnectivity graphs obtained after exhaustive runs on (L-Ala)$_6$ in vacuum and water (generalised Born approximation, implicit solvent) are depicted in Figures 4 and 5, respectively, with the helicity order parameter derived from both QS and CD methods. It has previously been shown for the (L-Ala)$_{12}$ isomer that helicity increases with increasing dielectric constant.[55,58] Our results show that the same can also be said of partial helicity for a short peptide of up to 6 residues. Similar behaviour is also observed for the (L-Ala-D-Ala)$_3$ system that does not form a completely helical structure as a low-energy isomer. It is interesting to note that we find more than one structure folded into a full $\alpha$-helix. While the global minimum is the best energetically, we find $\alpha$-helices of other radii to be present in the vicinity of the global minimum as well.

A comparison between fractional helicities calculated from CD and QS methods is depicted in Figure 6. QS captures more fully helical structures than the corresponding maximum

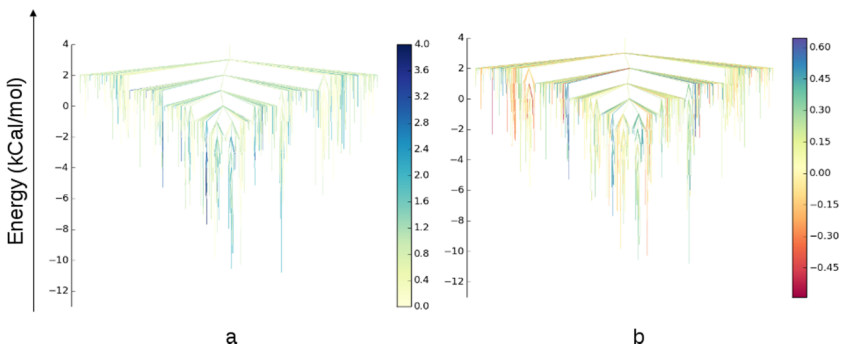


FIG. 4. Disconnectivity graphs for Ace-(L-Ala)$_6$-NHMe in vacuum coloured according to fractional helicity derived from (a) quaternion superimposition method and (b) circular dichroism. In (a), blue indicates complete helicity and yellow indicates none. In (b), yellow indicates no fractional helicity, blue indicates fractional helicity in the left sense, and red indicates fractional helicity in the right sense of the helix. The lowest 5000 minima are considered.

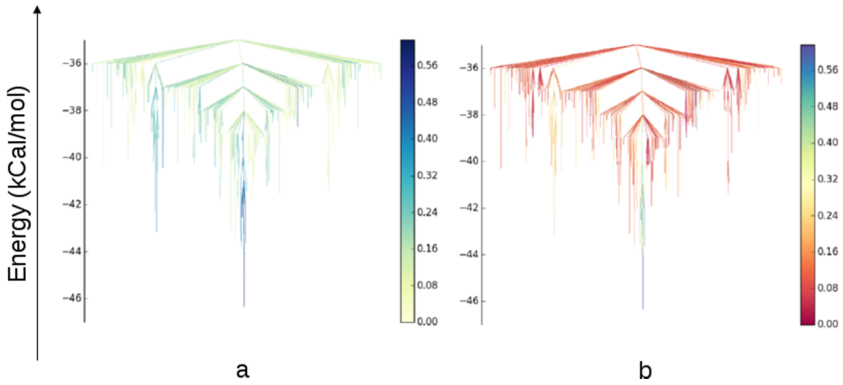


FIG. 5. Disconnectivity graphs for Ace-(L-Ala)$_6$-NHMe in water coloured according to fractional helicity derived from (a) quaternion superimposition method and (b) circular dichroism. The lowest 5000 minima are considered. Colour scheme is identical to Figure 4.

helicity measure from CD spectra. However, there is a correlation between the helical content per minimum calculated using both of these methods.

## Effects of changing backbone stereochemistry

### *Syndiotactic isomer*

Polymer chains with all stereocenters of exclusively L or D configurational type are called isotactic, homotactic, or homochiral polymers.[59] Chains of four amino acids can form one turn of an $\alpha$-helix, and the largest conformational winnowing effects were found in the $\alpha$-helical region.[60] Therefore, the low likelihood of syndiotactic sequences (Ace-(L-Ala-D-Ala)$_3$-NHMe) to form an $\alpha$-helix increases the total number of accessible conformations for short peptides. Larger syndiotactic sequences have been reported to fold in the $\alpha$-helical structure. We do not see any evidence of this for the hexaalanine system. From the disconnectivity graph, it is clear that a much greater configurational space is accessible to the syndiotactic system and the energy landscape of the isotactic polymer is much more funnelled. Qualitatively, the system becomes much more frustrated and glass-like for the alternating -(D,L)- system. The disconnectivity graphs along with the global minima are depicted in Figure 7. Here, we study only Ace-(D-Ala-L-Ala)$_3$-NHMe. We note that the landscape of Ace-(L-Ala-D-Ala)$_3$-NHMe system is expected to be a mirror image of this system and was studied for global optimisation runs in the subsection titled "Global optimisation with fixed chiralities".

The chart is coloured according to the fractional helicities derived from the calculated CD spectra. This measure of helicity is chosen particularly as, in addition to the magnitude of partial helicity, the sense of the helix is also indicated. As the chiralities are flipped, this information is vital. As in the case of the isotactic all-L polymers, the order parameters derived from quaternion superimposition, cluster compare, and CD spectra agree rather well. We can therefore state with confidence that the partial helicities are indeed reproducible from all the three methods mentioned above. In fact, we are able to locate on the disconnectivity chart the most helical right-handed and left-handed structures found for the isotactic polymer. The global minimum has a DLD sub-unit folded partly into a helical state where through-space interactions are between i, i + 5 residues.

Further, the most right-handed and the most-left-handed helical structures are separated by an energy of 2.714 kcal/mol and appear to belong to the same energetic basin. Linderstrøm-Lang synthesised poly-alanine with alternating D and L amino acids of 30 residues that form an $\alpha$-helix. It is possible that for longer chain lengths, these helical structures are further stabilized.

It has been suggested qualitatively that the protein homochirality facilitates the formation of $\alpha$-helices by funnelling the landscape significantly in favour of helical and partially

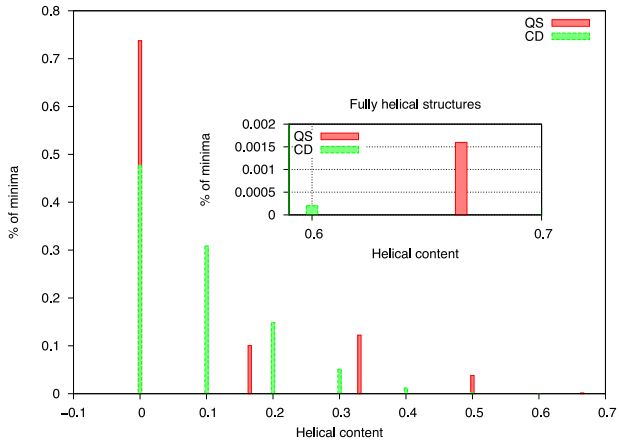


FIG. 6. Comparison of fractional helicities calculate from quaternion superimposition (QS) and CD spectra for each local minimum found in the database. Inset: Very low percentage of structures are characterised as fully helical.

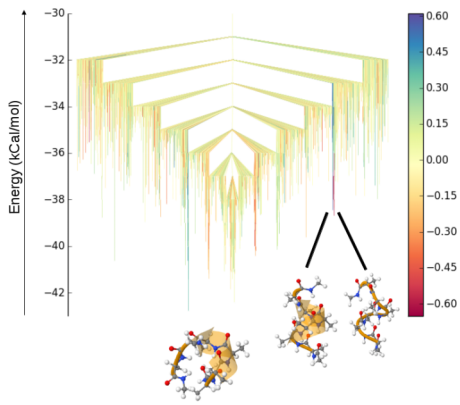


FIG. 7. Disconnectivity graph for Ace-(D-Ala-L-Ala)$_3$-NHMe in water coloured according to fractional helicities derived from calculated circular dichroism spectra. Colour scheme identical to Figure 4(b). The lowest energy 5000 minima are plotted.

TABLE II. Average fractional helicities for the ensemble derived from a circular dichroism from the database of minima for three stereoisomers of hexa-alanine.

| System | $\langle f_H \rangle$ ($\times 10^{-2}$) at 273 K | From other studies ($\times 10^{-2}$) at 273 K |
|---|---|---|
| LLLLLL | 8.044 | 2.848[49] and 4.297[47] |
| DLDLDL | +5.373 78 | |
| DDDLLL | −2.549 211 | |

helical structures.[11] While this can be observed clearly from the average values of fractional helicities weighted by energies presented in Table II, we must also note that fully helical structures of the syndiotactic polymer are in fact present as local minima and are probably further stabilized on lengthening the chain. The global minimum also comprises an LDL sequence folded into a distorted $\alpha$-helical configuration. This global minimum obtained from all-atom simulations with the AMBERff03 force-field is in fact different from the one obtained from lattice chain models by Nanda *et al.* and has no cyclic character as was inferred therein. This can be attributed to the force-field optimizing a combination of torsion angles between residues and the number of hydrogen bonds. The ability of syndiotactic hexa-alanine to form pseudo-cyclic peptides is evident from the behaviour of the syndiotactic zwitterionic structure. Here, we see that the dielectric constant of water is not sufficiently strong to screen the charged ends of the peptide and it invariably forms a pseudo-cyclic peptide. The next step is to observe the disconnectivity graphs and compare the fractional helicities derived for the Ace-(D-Ala)$_3$-(L-Ala)$_3$-NHMe hexamer (see Fig. 8) as this allows for (i, i + 3) hydrogen bonding that facilitates completion of one turn of an $\alpha$-helix. To our knowledge, the energy landscape of this system has not been explored in detail.

### Application to helix-coil transition theory

We are able to assign fractional helicity to every structure corresponding to a potential energy minimum from the aforementioned methods. This makes application of helix-coil transition model of Lifson-Roig[46] possible. Fractional helicities from CD spectra calculations are derived from Equation (1). We get a continuous distribution of fractional helicity values. As shown in Figure 6, the values from both QS and CD methods correlate rather well. Thus, the average fractional helicity $\langle f_H \rangle$ is derived from $f_H$ for each local minimum (≈12 000 local minima per stereoisomer in the database). Here, we use the results obtained from the CD calculations. A Boltzmann weighting is attached to the energies. The helix nucleation and propagation parameters for hexaalanine in water have been previously derived through experimental and theoretical means.[47,49,61] For a short chain with identical residues, within the one-sequence approximation, $\langle f_H \rangle$ is calculated from Equation (2.48) in Ref. 45.

We can then use our database of local minima to calculate average fractional helical content and the average length of the helix. Given a partition function, these data should give us access to the helix nucleation and helix propagation parameters within the Lifson-Roig helix-coil model for different stereoisomers of alanine. These are tabulated in Table II.

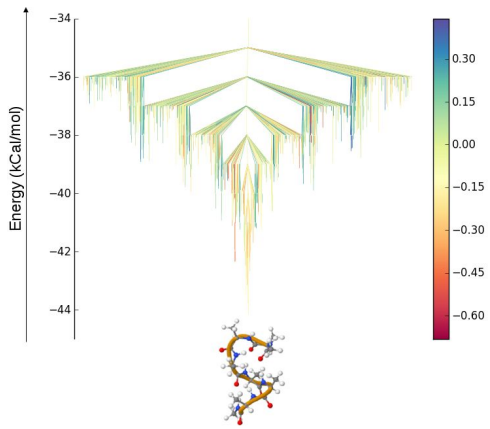


FIG. 8. Disconnectivity graph for Ace-(D-Ala)$_3$-(L-Ala)$_3$-NHMe in water coloured according to fractional helicities derived from calculated circular dichroism spectra. Colour scheme identical to Figure 4(b). The global minimum is depicted along with the metric disconnectivity graph. The lowest 5000 minima are plotted.

## CONCLUSIONS

We have studied the relationship between chirality and backbone structure for hexa-alanine by mapping the energy landscape of its stereoisomers. Energetically favourable global minima were determined through global optimisation methods. A database of connected low-energy local minima and transition points was constructed for stereoisomers and zwitterionic stereoisomers of hexa-alanine. In particular, three order parameters are described for calculating fractional helicity, and the effect of stereoisomerism on partial helicity was described by the use of metric disconnectivity graphs. Further, the Zimm-Bragg model for helix-coil transitions was used to derive Boltzmann weighted average of partial helicities and compared with previous studies.

## ACKNOWLEDGMENTS

We thank Professor Jonathan Hirst for providing the CD calculation software and acknowledge the Engineering and Physical Sciences Research Council, UK (EPSRC) for funding under Programme Grant No. EP/I001352/1. The computations described in this paper were performed using the University of Birmingham's BlueBEAR HPC service, which provides a High Performance Computing service to the University's research community. See http://www.birmingham.ac.uk/bear for more details.

[1]L. Pauling and R. Corey, Proc. Natl. Acad. Sci. U. S. A. **37**, 241–250 (1951).
[2]G. Wald, Ann. N. Y. Acad. Sci. **69**, 352–368 (1957).
[3]J. F. Yan, F. A. Momany, and H. A. Scheraga, J. Am. Chem. Soc. **92**, 1109–1115 (1970).
[4]T. Ooi, R. A. Scott, G. Vanderkooi, and H. A. Scheraga, J. Chem. Phys. **46**, 4410–4426 (1967).
[5]S. Martínez-Rodríguez, A. I. Martínez-Gómez, F. Rodríguez-Vico, J. M. Clemente-Jiménez, and F. J. Las Heras-Vázquez, Chem. Biodiversity **7**, 1531–1548 (2010).
[6]S. A. Fuchs, R. Berger, L. W. J. Klomp, and T. J. de Koning, Mol. Genet. Metab. **85**, 168–180 (2005).

[7]S. Rana, B. Kundu, and S. Durani, Chem. Commun. **2005**, 207–209.
[8]V. Ramakrishnan, R. Ranbhor, and S. Durani, Biopolymers **78**, 96–105 (2005).
[9]G. Schmidt, D. C. Hodgkin, and B. M. Oughton, Biochem. J. **65**, 744 (1957).
[10]S. Hovmöller, T. Zhou, and T. Ohlson, Acta Crystallogr., Sect. D: Biol. Crystallogr. **58**, 768–776 (2002).
[11]V. Nanda, A. Andrianarijaona, and C. Narayanan, Protein Sci. **16**, 1667–1675 (2007).
[12]W. Kabsch and C. Sander, Biopolymers **22**, 2577–2637 (1983).
[13]V. Nanda and W. F. DeGrado, J. Am. Chem. Soc. **126**, 14459–14467 (2004).
[14]J. Hermans, A. G. Anderson, and R. H. Yun, Biochemistry **31**, 5646–5653 (1992).
[15]D. C. Poland and H. A. Scheraga, J. Chem. Phys. **43**, 2071–2074 (1965).
[16]R. Frühwirth, A. Strandlie, and W. Waltenberger, Nucl. Instrum. Methods Phys. Res., Sect. A **490**, 366–378 (2002).
[17]J. Christopher, R. Swanson, and T. Baldwin, Comput. Chem. **20**, 339–345 (1996).
[18]Y. Nievergelt, Comput. Aided Geom. Des. **14**, 707–718 (1997).
[19]P. Enkhbayar and N. Matsushima, AIP Conf. Proc. **1479**, 1942–1947 (2012).
[20]P. Kumar and M. Bansal, J. Biomol. Struct. Dyn. **30**, 773–783 (2012).
[21]O. M. Becker and M. Karplus, J. Chem. Phys. **106**, 1495–1517 (1997).
[22]D. J. Wales, M. A. Miller, and T. R. Walsh, Nature **394**, 758–760 (1998).
[23]L. C. Smeeton, M. T. Oakley, and R. L. Johnston, J. Comput. Chem. **35**, 1481–1490 (2014).
[24]D. A. Case *et al.*, AMBER 12, 2012, see http://ambermd.org/.
[25]M. C. Lee and Y. Duan, Proteins: Struct., Funct., Bioinf. **55**, 620–634 (2004).
[26]A. Onufriev, D. Bashford, and D. A. Case, Proteins: Struct., Funct., Bioinf. **55**, 383–394 (2004).
[27]D. J. Wales and J. P. K. Doye, J. Phys. Chem. A **101**, 5111–5116 (1998).
[28]D. J. Wales, GMIN: A program for finding global minima and calculating thermodynamic properties from basin-sampling, see http://www-wales.ch.cam.ac.uk/GMIN/; last accessed 20/10/2015.
[29]D. J. Wales, PATHSAMPLE: A program for refining and analysing kinetic transition networks, see http://www-wales.ch.cam.ac.uk/PATHSAMPLE/; last accessed 20/10/2015.
[30]J. M. Carr, S. A. Trygubenko, and D. J. Wales, J. Chem. Phys. **122**, 234903 (2005).
[31]S. A. Trygubenko and D. J. Wales, J. Chem. Phys. **120**, 2082–2094 (2004).
[32]M. S. Bauer, B. Strodel, S. N. Fejer, E. F. Koslover, and D. J. Wales, J. Chem. Phys. **132**, 054101 (2010).
[33]D. J. Wales, OPTIM: A program for characterising stationary points and reaction pathways, see http://www-wales.ch.cam.ac.uk/OPTIM/; last accessed 20/10/2015.
[34]J. Aqvist, Comput. Chem. **10**, 97–99 (1986).
[35]R. Hundt, J. C. Schön, S. Neelamraju, J. Zagorac, and M. Jansen, J. Appl. Crystallogr. **46**, 587–593 (2013).
[36]G. R. Kneller and P. Calligari, Acta Crystallogr., Sect. D: Biol. Crystallogr. **62**, 302–311 (2006).
[37]B. M. Bulheller, A. Rodger, and J. D. Hirst, Phys. Chem. Chem. Phys. **9**, 2020–2035 (2007).
[38]R. W. Woody, Jr., J. Chem. Phys. **46**, 4927–4945 (1967).
[39]R. W. Woody, J. Chem. Phys. **49**, 4797–4806 (1968).
[40]P. M. Bayley, E. B. Nielsen, and J. A. Schellman, J. Phys. Chem. **73**, 228–243 (1969).
[41]J. D. Hirst, K. Colella, and A. T. B. Gilbert, J. Phys. Chem. B **107**, 11813–11819 (2003).
[42]M. T. Oakley, B. M. Bulheller, and J. D. Hirst, Chirality **18**, 340–347 (2006).
[43]B. M. Bulheller and J. D. Hirst, Bioinformatics **25**, 539–540 (2009).
[44]N. A. Besley and J. D. Hirst, J. Am. Chem. Soc. **121**, 9636–9644 (1999).
[45]D. Poland and H. Scheraga, *Theory of Helix-Coil Transitions in Biopolymers: Statistical Mechanical Theory of Order-Disorder Transitions in Biological Macromolecules*, Molecular Biology (Academic Press, New York, 1970).
[46]S. Lifson and A. Roig, J. Chem. Phys. **34**, 1963 (1961).
[47]S. Gnanakaran and A. E. García, Proteins: Struct., Funct., Bioinf. **59**, 773–782 (2005).
[48]A. Chakrabartty, J. A. Schellman, and R. L. Baldwin, Nature **351**, 586–588 (1991).
[49]J. M. Scholtz and R. L. Baldwin, Annu. Rev. Biophys. Biomol. Struct. **21**, 95–118 (1992).
[50]J. Wójcik, K.-H. Altmann, and H. A. Scheraga, Biopolymers **30**, 121–134 (1990).
[51]H. A. Scheraga, Pure Appl. Chem. **36**, 1–8 (1973).
[52]There cannot be more than one set of helices within the chain. For a six-residue peptide, this is a reasonable assumption to make as one complete turn of an alpha-helix requires at least 3 consecutive residues to be in a helical state.
[53]C. Nick Pace and J. Martin Scholtz, Biophys. J. **75**, 422–427 (1998).
[54]S. Somani and D. J. Wales, J. Chem. Phys. **139**, 121909 (2013).
[55]P. N. Mortenson, D. A. Evans, and D. J. Wales, J. Chem. Phys. **117**, 1363 (2002).
[56]A. Berger and K. Linderstrøm-Lang, Arch. Biochem. Biophys. **69**, 106–118 (1957).
[57]N. Lotan, F. T. Hesselink, H. Benderly, J. Yan, I. Schechter, A. Berger, and H. Scheraga, Macromolecules **6**, 447–453 (1973).
[58]Y. Levy, J. Jortner, and O. M. Becker, Proc. Natl. Acad. Sci. U. S. A. **98**, 2188–2193 (2001).
[59]F. W. Billmeyer, *Textbook of Polymer Science* (Wiley, 1984).
[60]R. V. Pappu, R. Srinivasan, and G. D. Rose, Proc. Natl. Acad. Sci. U. S. A. **97**, 12565–12570 (2000).
[61]K. Song, J. M. Stewart, R. M. Fesinmeyer, N. H. Andersen, and C. Simmerling, Biopolymers **89**, 747–760 (2008).